

us-10-038-854-38

GenCore version 5.1.6
copyright (c). 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 29, 2004, 21:29:29 ; Search time 2766 Seconds
(without alignments)

5556.365 Million cell updates/sec

Title: US-10-038-854-38

Perfect score: 14495

Sequence: 1 MDVKERRPYCSDLTKSREKE.....ELADSANNIQFLRQSEIGRR 2721

Scoring table: BLOSUM62

Xgapop	10.0	,	Xgapext	0.5
Ygapop	10.0	,	Ygapext	0.5
Fgapop	6.0	,	Fgapext	7.0
Delop	6.0	,	Delext	7.0

Searched: 4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters: 8352472

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+p2n.model -DEV=x1h
-Q=/cgn2_1/uspto_spool/us10038854/runat_29122004_112455_10566/app_query.fasta_1.2887
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10038854@CGN_1_1_2038@runat_29122004_112455_10566
```

us-10-038-854-38
 -NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WATT -DSPBLOCK=100
 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*

```

1: /cgn2_6/ptodata/2/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/us06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/us06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/us07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/us08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/us08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/us09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/us09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/us09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/us09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/us10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/us10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/us10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/us10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/us10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/us10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/us11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/us60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/us60_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14495	100.0	8645	16	US-10-038-854-37	Sequence 37, Appl
2	14475	99.9	8675	16	US-10-038-854-35	Sequence 35, Appl
3	13944.5	96.2	8473	16	US-10-038-854-39	Sequence 39, Appl
4	13834	95.4	8487	16	US-10-038-854-41	Sequence 41, Appl

3

us-10-038-854-38

5	10403	71.8	8689	9	US-09-808-602-78	Sequence 78, Appl
6	10403	71.8	8689	10	US-09-800-198-66	Sequence 66, Appl
7	10394.5	71.7	8409	9	US-09-808-602-79	Sequence 79, Appl
8	10394.5	71.7	8409	10	US-09-800-198-67	Sequence 67, Appl
9	10394	71.7	8575	16	US-10-072-012-143	Sequence 143, App
10	10393.5	71.7	8797	9	US-09-808-602-74	Sequence 74, Appl
11	10393.5	71.7	8797	9	US-09-808-602-77	Sequence 77, Appl
12	10393.5	71.7	8797	10	US-09-800-198-62	Sequence 62, Appl
13	10393.5	71.7	8797	10	US-09-800-198-65	Sequence 65, Appl
14	10329.5	71.3	8438	16	US-10-042-865-1	Sequence 1, Appl
15	10310.5	71.1	9729	9	US-09-808-602-12	Sequence 12, Appl
16	10310.5	71.1	9729	10	US-09-800-198-12	Sequence 12, Appl
17	10309	71.1	9826	9	US-09-808-602-7	Sequence 7, Appl
18	10309	71.1	9826	10	US-09-800-198-7	Sequence 7, Appl
19	10221	70.5	8354	16	US-10-383-201-43	Sequence 43, Appl
20	10221	70.5	8354	16	US-10-029-020-13	Sequence 13, Appl
21	10078	69.5	9695	15	US-10-144-194A-81	Sequence 81, Appl
22	10078	69.5	9695	18	US-10-491-566-81	Sequence 81, Appl
23	10010	69.1	8355	16	US-10-383-201-55	Sequence 55, Appl
24	9765	67.4	6810	18	US-10-723-860-8301	Sequence 8301, Ap
25	9491	65.5	9058	15	US-10-144-194A-79	Sequence 79, Appl
26	9491	65.5	9058	18	US-10-491-566-79	Sequence 79, Appl
27	9267	63.9	5309	18	US-10-723-860-4493	Sequence 4493, Ap
28	8884	61.3	8297	18	US-10-723-860-4101	Sequence 4101, Ap
29	8884	61.3	12880	15	US-10-295-027-927	Sequence 927, App
30	8463.5	58.4	7781	18	US-10-723-860-2302	Sequence 2302, Ap
31	6894.5	47.6	6771	18	US-10-723-860-6509	Sequence 6509, Ap
32	6702.5	46.2	6560	9	US-09-808-602-76	Sequence 76, Appl
33	6702.5	46.2	6560	10	US-09-800-198-64	Sequence 64, Appl
34	5714	39.4	3614	15	US-10-172-118-1743	Sequence 1743, Ap
35	5714	39.4	3614	16	US-10-342-887-1743	Sequence 1743, Ap
36	4678	32.3	8624	18	US-10-723-860-3169	Sequence 3169, Ap
37	4577.5	31.6	8774	18	US-10-723-860-7176	Sequence 7176, Ap
38	2784	19.2	2496	9	US-09-808-602-75	Sequence 75, Appl
39	2784	19.2	2496	10	US-09-800-198-63	Sequence 63, Appl
40	2316	16.0	1534	18	US-10-128-558-121	Sequence 121, App
41	2263	15.6	3217	14	US-10-198-846-13976	Sequence 13976, A
42	2002	13.8	1399	16	US-10-383-201-49	Sequence 49, Appl
43	2002	13.8	1399	16	US-10-383-201-59	Sequence 59, Appl
44	1954.5	13.5	1371	16	US-10-383-201-51	Sequence 51, Appl

45 1824 12.6 1476 16 US-10-383-201-41 us-10-038-854-38

Sequence 41, Appl

ALIGNMENTS

RESULT 1

US-10-038-854-37

Sequence 37, Application US/10038854

Publication No. US20040022781A1

GENERAL INFORMATION:

APPLICANT: Spytek, Kimberly A
APPLICANT: Li, Li
APPLICANT: Wolenc, Adam R
APPLICANT: Vernet, Corine
APPLICANT: Eisen, Andrew J
APPLICANT: Liu, Xiaohong
APPLICANT: Malyankar, Uriel M
APPLICANT: Shimkets, Richard A
APPLICANT: Tchernev, Velizar
APPLICANT: Spaderna, Steven K
APPLICANT: Gorman, Linda
APPLICANT: Kekuda, Ramesh
APPLICANT: Patturajan, Meera
APPLICANT: Gusev, Vladimir Y
APPLICANT: Gangolli, Esha A
APPLICANT: Guo, Xiaojaia S
APPLICANT: Shenoy, Suresh G
APPLICANT: Rastelli, Luca
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc
APPLICANT: Burgess, Catherine E
APPLICANT: Edinger, Shlomit R
APPLICANT: Ellerman, Karen
APPLICANT: Gunther, Erik
APPLICANT: Smithson, Glenna
APPLICANT: Millet, Isabelle
APPLICANT: MacDougall, John R
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-230
CURRENT APPLICATION NUMBER: US/10/038,854

us-10-038-854-38

; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258, 928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259, 415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259, 785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269, 814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279, 832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279, 833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279, 863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283, 889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284, 447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286, 683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-038-854-37

Alignment Scores:
Pred. No.: 0
Score: 14495.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 16

Length: 8645
Matches: 2721
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-038-854-38 (1-2721) x US-10-038-854-37 (1-8645)

us-10-038-854-38

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:

December 29, 2004, 17:19:17 ; Search time 23977 Seconds
(without alignments)
5366.613 Million cell updates/sec

Title: US-10-038-854-38

Perfect score: 14495

Sequence: 1 MDVKERRPYCSLTKSREKE. ELADSANNIQFLRQSEIGRR 2721

Scoring table:

BLOSUM62			
Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US10038854/runat_29122004_112453_10542/app_query.fasta_1.2887
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pfo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10038854@CGN_1_1_15641@runat_29122004_112453_10542 -NCPU=6 -ICPU=3
```

us-10-038-854-38

-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:*2: gb_htg:*3: gb_in:*4: gb_om:*5: gb_ov:*6: gb_pat:*7: gb_ph:*8: gb_pl:*9: gb_pr:*10: gb_ro:*11: gb_sts:*12: gb_sy:*13: gb_un:*14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14495	100.0	8645	6 AX662355	AX662355 Sequence
2	14475	99.9	8675	6 AX662353	AX662353 Sequence
3	14248	98.3	8964	6 AX952856	AX952856 Sequence
4	14248	98.3	8964	10 AB025412	AB025412 Mus muscu
5	13944.5	96.2	8473	6 AX662357	AX662357 Sequence
6	13834	95.4	8487	6 AX662359	AX662359 Sequence
7	12351	85.2	7816	10 AF195418	AF195418 Mus muscu
8	11770.5	81.2	6751	6 CQ716754	CQ716754 Sequence
9	11633	80.3	8816	5 AB026979	AB026979 Danio rer
10	10403	71.8	8689	6 AX250067	AX250067 Sequence
11	10403	71.8	8689	10 AF086607	AF086607 Rattus no

us-10-038-854-38
 12 10394.5 71.7 8409 5 GGA279031 AJ279031 Gallus ga.
 13 10394.5 71.7 8409 6 AX250068 Sequence
 14 10394 71.7 8575 6 AX921803 Sequence
 15 10393.5 71.7 8797 6 CQ777485 Sequence
 16 10393.5 71.7 8797 6 AX250063 Sequence
 17 10393.5 71.7 8797 6 AX250066 Sequence
 18 10393.5 71.7 8797 10 AB025411 AB025411 Mus muscu
 19 10329.5 71.3 8438 6 AX675551 Sequence
 20 10310.5 71.1 9729 6 AX250013 Sequence
 21 10309 71.1 9826 6 AX250008 Sequence
 22 10221 70.5 8354 6 AX556500 Sequence
 23 10189 70.3 8585 10 AB025413 AB025413 Mus muscu
 24 10144.5 70.0 8645 6 AX600210 Sequence
 25 10001 69.0 9722 10 AF059485 AF059485 Mus muscu
 26 9764.5 67.4 9264 5 AB026980 AB026980 Danio rer
 27 9489 65.5 5804 10 AK122513 AK122513 Mus muscu
 28 9432 65.1 7400 6 CQ727408 Sequence
 29 9267 63.9 5309 9 AB040888 AB040888 Homo sapi
 30 9194 63.4 7514 6 CQ722991 CQ722991 Sequence
 31 9035 62.3 8118 5 GGA238613 AJ238613 Gallus ga
 32 8993.5 62.0 7713 9 HSM808325 BX648178 Homo sapi
 33 8972 61.9 8373 10 AB025410 AB025410 Mus muscu
 34 8884 61.3 8297 9 AF100772 AF100772 Homo sapi
 35 8663 59.8 7706 6 CQ714850 CQ714850 Sequence
 36 8463.5 58.4 7781 9 AB032953 AB032953 Homo sapi
 37 6702.5 46.2 6560 6 AX250065 AX250065 Sequence
 38 5955.5 41.1 3394 9 AK125869 AK125869 Homo sapi
 39 5714 39.4 3614 6 AX876525 AX876525 Sequence
 40 5714 39.4 3614 6 BD156175 BD156175 Primer fo
 41 5714 39.4 3614 9 AK001336 AK001336 Homo sapi
 42 5313 36.7 8993 9 HSM806812 BX640737 Homo sapi
 43 5117 35.3 3270 6 AX877449 AX877449 Sequence
 44 5117 35.3 3270 6 BD156663 BD156663 Primer fo
 45 5117 35.3 3270 9 AK001748 AK001748 Homo sapi

Search completed: December 30, 2004, 05:27:18
 Job time : 25852 secs